Questions and Answers from Listening Session: Complication and Comorbidity (CC), Major Complication and Comorbidity Comprehensive (MCC) Severity Level Designation- October 8, 2019

1. So my first question is, how do you account for common resource utilization? So if a patient has an acute exacerbation of COPD and they have acute and chronic respiratory failure, they may have common resource utilization. How can you tease that out and see whether the impact is what it would be for some – like I can understand if you have a CC that's completely unrelated then you can see how this particularly affects the resource utilization, but if they have common resource utilization, I don't understand how you could necessarily determine that this – you know it's not necessarily additive, so I don't understand that piece of it.
    a. It's not exactly additive. You're trying to get at a designation of either CC, non-CC, or MCC. You're not trying to pull out a regression formula with a coefficient value, and that's where that would be a whole separate conversation, and then you have a lot of other challenges with that approach. So what we're trying to do is create a way to designate the level, and then you re-compute as an iterative process, you would then reestablish all of these variables and all of these statistics to see, with a new system and the diagnosis codes falling into their new designation level, if that diagnosis and the other diagnosis are then meeting the expectations or the expected values. And at the end of the day, again, it's a clinical decision, so the statistics may not fully tease out ever. Ultimately when there are differences in the evaluation of the statistics and the clinical approach, the teams typically go with the clinical designation, and just simply use the statistics as a way to kind of validate aspects.

2. Is it possible to every year report these thresholds you're using for the different levels? Second is, can you on this tab with-- one where you show the different impacts, is it possible to show the raw dollars that have been computed for each code? That's sometimes useful to see for those of us who are trying to replicate. And the third question is more a request, if you could describe a little bit of, where does the CC exclusion list come in and how is that developed?
    a. Thank you for those comments. We would be happy to take your request into consideration. We're considering those types of things, as well as making a (workbook) publicly available for in the future. As far as this report is concerned, if you're asking about the CC exclusion list and how that applies to this report, any secondary diagnosis code that is excluded due to the principal diagnosis would not be evaluated in this report, it would be excluded.

3. The question we struggled with during the comment period had to do with the interaction-- in other words, when you decide to downgrade a CC or, you know, make it no longer a CC or a MCC, then those cases change subclasses, and how do you consider the interactive effect of dropping certain codes as you evaluate each subsequent code? Do you know what I mean? In other words, something that looks like it's not different from others in the subclass might look different if the subclass changes.
    a. Absolutely. And as part of this process, we don't just go through this once, we'll then simulate the impact of making these changes. Once each of the diagnosis codes were evaluated, the ones that were revised, these computations are re-estimated and the computational values for each secondary diagnosis code that had changes with the list, so you know that where on the CC or major CC changes were made, they were all revised and then reevaluated, relooked at. You know, we went through a lot of effort to document aspects of statistically met criteria, clinically made a decision, all of that, so when it's redone and reevaluated and re-reviewed, we would look especially at those that changed to make sure that nothing else either popped up because other secondary diagnosis codes moved, and then therefore it would move groups of cases from one subgroup to another. The expected's would change, and all of that would get reevaluated. So it's an iterative process until it was stable.
        1. In general, how many iterations is required?
            a. It really just depends on the codes. It really just depends.
        2. Any additional information you could provide about that, you know, in terms of letting us see different versions of it would be very helpful. My second question, though, is just doing a quick and dirty, you know, analysis of the new table that you provided, it does seem like you're providing the outcome for all of the codes with at least 200 cases – well actually no there's – for all the codes period, but that's one of the things that we ask for. Is that correct?
            a. Yes, that's correct.
        3. Okay, and then it looks like the number of cases with a suggested change is much lower. Do you agree with that, and also that there are fewer downgrades or more upgrades? Can you just comment, you know, on where your current thinking is vis-á-vis the proposed rule?
            a. So, in preparation for this listening session, we took into consideration some of the public feedback that we had gotten requesting this additional information, and so we did, you know, put out the additional information to help explain the methodology, and to give folks a better sense of how we use it in our decision making process. But part of – the purpose of this listening session was to help us get a little bit more feedback from all of the public and the stakeholders in order to prepare for the 2021 rulemaking.

And so at this time, we really haven't started revisiting that process yet. We're still collecting and gathering information.

4. So we shouldn't interpret the information in this file as suggesting anything for the 2021 proposed rule?
    a. That's correct.

4. I have a question with regard to the flag criteria and the volume of cases, because like the prior comments are – I'm looking in your second tab, the 1-SDX Codes, I see volume is very low, but I do see what appears to be a recommended status change from a category. And so when you talked about the family of codes, did you ever consider aggregating across the family, like the neoplasms, for example, and seeing if the results changed?
    a. We did not aggregate across family of codes. The clinical teams upon their review would look at the whole family of codes, put together their principles from which-- and discuss clinically the aspects of these secondary diagnosis codes and the impact they are having, and then ultimately make decisions to change a designation or not. The flagging criteria again is purely statistical. No change simply means that it didn't hit the flagging criteria to recommend an increase or a decrease. It could be because there was just not enough volume, low volume, so they all just said "no change." We could probably improve the flagging criteria to get out of – to put LV for low volume or something else so that "no change" truly recommended that statistically there was volume, but statistically it is good with the current designation so that's probably an improvement that we can make. So I just want to make sure everyone is aware of that. The other thing that I just want to also clarify is that on the report you'll see in Column C and D, if the diagnosis code is new, it has a N. Obviously there's not going to be any data to support it, but the old deleted code is probably in the data. So we will also then reference for the new code what the deleted mapped code is so that the clinical teams can then look back at that mapped code, look at those statistics and then ultimately make a decision as to how to apply that information, those statistics, forward to the new codes, and decide clinically what they want to do with the designation.

5. I couldn't understand from the Federal Register that the proposed note, how moderate malnutrition could have a higher category than severe malnutrition. Was that just an error or how could that be a result of an iterative clinical process?
    a. Thank you for that question. That was similar to some of the comments that we received on the proposed rule. You know the purpose of this call was really just to sort of give a broad overview of the process and we can-- for any detailed comments, we encourage you to please put those in writing to our MS-DRG change request mailbox, and I'll be giving that mailbox at the end of the call and then we can take that feedback into consideration. Thank you.

6. I think I recall going through an exercise similar to this probably about 12 years ago for the MS-DRGs. I see on Tab 1-SDX codes there are 64,884 lines in that spreadsheet. So I presume what you would do is only evaluate and provide clinical oversight to the ones where there is a change or recommendation going up or going down?
    - **a.** So the clinical team would take an evaluation not only of those that showed a flag indicated statistically, but also would look down and clinically look at family of codes to do that analysis as well. And so, yes, it's a lot of work to go through individual-- code by code by code-- and for the most part, the clinical teams --the family of codes – the broad set of codes, but absolutely they would evaluate every single code that had a flag indicating statistically to make a change, but those are not the only codes that were looked at. Again, a broad family of codes, to make sure there was consistency as well as they could across the codes, and the principles around making those decisions.
        1. This is the first time this is being done with ICD-10? So the current classifications are based on the crosswalks between ICD-9 and ICD-10?
            a. That is correct.
7. In your methodology, on the flag criteria tab, it depends on the volume being greater than 200. We have greater granularity in the ICD-10 than ICD-9. Did you adjust the threshold when moving to ICD-10, or did everything stay the same? Just wondering if some things might not be picked up because of the greater granularity.
    - **a.** So when this research was done back in 2007-- and yes, I was a participant in those meetings as well-- we really didn't have official flagging criteria. It was literally large, long meetings with rooms of clinicians going through every single code row by row, one by one, looking at the family of codes, looking at all of the information. There was no hard, fast kind of suggested rule criteria at that moment. So what the analysts did who have been working on these types of projects for the last 40 years-- not myself, only the last 30 years for myself-- we came up with this general flagging criteria. It's kind of like the white flag, it's not the red flag, stop, do nothing, don't you know – stop, don't pass go kind of thing. It's the white flag that says, hey, this is something to look at, and the teams kind of came up with around 200 cases was the right amount for this type of analysis – again, it's just a flag. The teams look at things, they step back, they look at the family of codes, and then they look at everything around that to ultimately make their decision.